

ENTROPY, STOCHASTIC MATRICES, AND QUANTUM OPERATIONS

LIN ZHANG

ABSTRACT. The goal of the present paper is to derive some conditions on saturation of (strong) subadditivity inequality for the stochastic matrices. The notion of relative entropy of stochastic matrices is introduced by mimicking quantum relative entropy. Some properties of this concept are listed and the connection between the entropy of the stochastic quantum operations and that of stochastic matrices are discussed.

1. INTRODUCTION

If the column vectors $\mathbf{p} = [p_1, \dots, p_N]^T \in \mathbb{R}^N$ and $\mathbf{q} = [q_1, \dots, q_N]^T \in \mathbb{R}^N$ are two probability distributions, the *Shannon entropy* of \mathbf{p} is defined by $H(\mathbf{p}) = -\sum_{i=1}^N p_i \log_2 p_i$, and the *relative entropy* of \mathbf{p} and \mathbf{q} is defined by $H(\mathbf{p}||\mathbf{q}) = \sum_{i=1}^N p_i \log_2 \frac{p_i}{q_i}$ when \mathbf{p} is absolute continuously with respect to \mathbf{q} , where $x \log_2 x$ is set to 0 if $x = 0$; $H(\mathbf{p}||\mathbf{q}) = +\infty$ otherwise.

Let $B = [b_{ij}]$ be a $N \times N$ bi-stochastic matrix, that is, $b_{ij} \geq 0$, and $\sum_{i=1}^N b_{ij} = \sum_{j=1}^N b_{ij} = 1$ for each $i, j = 1, \dots, N$. Let π be a permutation of the set $\{1, \dots, N\}$. For any $i, j \in \{1, \dots, N\}$, we define $c_{ij} = 1$ when $i = \pi(j)$ and $c_{ij} = 0$ when $i \neq \pi(j)$. Then the matrix $C = [c_{ij}]$ is called a permutation matrix. Let \mathbf{S}_N be the set of all $N \times N$ permutation matrices and \mathbf{B}_N be the convex hull \mathbf{B}_N of \mathbf{S}_N . The well-known Birkhoff-von Neumann theorem indicates that \mathbf{B}_N is the set of all $N \times N$ bi-stochastic matrix.

We only consider finite dimensional complex Hilbert spaces. A state ρ of quantum system, described by Hilbert space \mathcal{H} , is a positive semi-definite matrix of trace one and call it the *density matrix*. The set of all density matrices of \mathcal{H} is denoted by $\mathbf{D}(\mathcal{H})$, if $\rho \in \mathbf{D}(\mathcal{H})$ is invertible, then ρ is said to be *faithful*. If ρ and σ are two quantum states, then the *von Neumann entropy* of ρ is defined by $S(\rho) = -\text{Tr}(\rho \log_2 \rho)$, the quantum relative entropy between ρ and σ is defined by $S(\rho||\sigma) = \text{Tr}(\rho(\log_2 \rho - \log_2 \sigma))$ if $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$; $S(\rho||\sigma) = +\infty$ otherwise, see [7].

Let \mathcal{H} and \mathcal{K} be two Hilbert spaces, $\mathbf{L}(\mathcal{H}, \mathcal{K})$ be the set of all linear operators from \mathcal{H} to \mathcal{K} , denote $\mathbf{L}(\mathcal{H}, \mathcal{H})$ by $\mathbf{L}(\mathcal{H})$. Let $\mathbf{T}(\mathcal{H}, \mathcal{K})$ denote the set of all *linear super-operators* from $\mathbf{L}(\mathcal{H})$ to $\mathbf{L}(\mathcal{K})$, similarly, denote $\mathbf{T}(\mathcal{H}, \mathcal{H})$ by $\mathbf{T}(\mathcal{H})$. We say that $\Phi \in \mathbf{T}(\mathcal{H}, \mathcal{H})$ to be *completely positive* (CP) if for each $k \in \mathbb{N}$, $\Phi \otimes \mathbb{1}_{\mathbf{M}_k(\mathbb{C})} : \mathbf{L}(\mathcal{H}) \otimes \mathbf{M}_k(\mathbb{C}) \rightarrow \mathbf{L}(\mathcal{K}) \otimes \mathbf{M}_k(\mathbb{C})$ is positive, where $\mathbf{M}_k(\mathbb{C})$ is the set of all $k \times k$ complex matrices. It follows from the famous theorems of Choi [2] and Kraus [3] that Φ can be represented in the form $\Phi = \sum_j \mathbf{A}_j \mathbf{A}_{M_j}^\dagger$,

where $\{M_j\}_{j=1}^n \subseteq \mathbf{L}(\mathcal{H}, \mathcal{K})$, that is, $\Phi(X) = \sum_{j=1}^n M_j X M_j^\dagger$, $X \in \mathbf{L}(\mathcal{H})$. Throughout the present paper, \dagger means adjoint operation of some operator. Denote by $\mathbf{CP}(\mathcal{H}, \mathcal{K})$ ($\mathbf{CP}(\mathcal{H})$) the set of all linear CP super-operators in $\mathbf{T}(\mathcal{H}, \mathcal{K})$ ($\mathbf{T}(\mathcal{H})$).

The so-called *quantum operation* is just a trace non-increasing $\Phi \in \mathbf{CP}(\mathcal{H}, \mathcal{K})$, if Φ is trace-preserving, then it is called *stochastic*; if Φ is stochastic and unit-preserving, then it is called *bi-stochastic*.

The famous *Jamiołkowski isomorphism* $J : \mathbf{T}(\mathcal{H}) \longrightarrow \mathbf{L}(\mathcal{H} \otimes \mathcal{H})$ transforms each $\Phi \in \mathbf{T}(\mathcal{H})$ into an operator $J(\Phi) \in \mathbf{L}(\mathcal{H} \otimes \mathcal{H})$, where $J(\Phi) = \Phi \otimes \mathbb{1}_{\mathbf{L}(\mathcal{H})}(\text{vec}(\mathbb{1}_{\mathcal{H}}) \text{vec}(\mathbb{1}_{\mathcal{H}})^\dagger)$. If $\Phi \in \mathbf{CP}(\mathcal{H})$, then $J(\Phi)$ is a positive semi-definite operator, in particular, if Φ is stochastic, then $\frac{1}{N}J(\Phi)$ is a state on $\mathcal{H} \otimes \mathcal{H}$. If $\Phi \in \mathbf{CP}(\mathcal{H})$ is a stochastic quantum operation, we denote the von Neumann entropy $\mathcal{S}(\frac{1}{N}J(\Phi))$ of $\frac{1}{N}J(\Phi)$ by $\mathcal{S}^{\text{map}}(\Phi)$ and call it the *map entropy* [10], which describes the decoherence induced by the quantum operation Φ .

2. ON SATURATION OF CLASSICAL RELATIVE ENTROPY

In order to obtain the condition for saturation of classical relative entropy, we need the following lemmas.

Lemma 2.1. ([5]) *Let \mathcal{H} be a Hilbert space, ρ and σ be two states of \mathcal{H} . If $\Phi \in \mathbf{CP}(\mathcal{H})$ is stochastic, then $\mathcal{S}(\Phi(\rho) \parallel \Phi(\sigma)) \leq \mathcal{S}(\rho \parallel \sigma)$.*

Lemma 2.2. ([4]) *Let $\{A_1, \dots, A_k\} \subseteq \mathbf{L}(\mathbb{C}^n)$ and $\{B_1, \dots, B_k\} \subseteq \mathbf{L}(\mathbb{C}^m)$ be two commuting families of Hermitian matrices. Then there exist unitary matrices $U \in \mathbf{L}(\mathbb{C}^n)$ and $V \in \mathbf{L}(\mathbb{C}^m)$ such that $U^\dagger A_j U$ and $V B_j V^\dagger$ are diagonal matrices with diagonals $\mathbf{a}_j = [a_{1j}, \dots, a_{nj}]^\top$ and $\mathbf{b}_j = [b_{1j}, \dots, b_{mj}]^\top$, respectively, for $j = 1, \dots, k$. Then the following conditions are equivalent:*

- (1) *There is a super-operator $\Phi \in \mathbf{CP}(\mathbb{C}^n, \mathbb{C}^m)$ such that $\Phi(A_j) = B_j$ ($j = 1, \dots, k$).*
- (2) *There is an $m \times n$ non-negative matrix $D = [d_{\mu\nu}]$ such that $[b_{ij}] = D[a_{ij}]$.*

Moreover, if the statement (2) is satisfied, then Φ is bi-stochastic if and only if D is bi-stochastic.

Theorem 2.3. *Let T be a $N \times N$ stochastic matrix, $\mathbf{p} = [p_1, p_2, \dots, p_N]^\top$ and $\mathbf{q} = [q_1, q_2, \dots, q_N]^\top$ be two N -dimensional probability distributions. Then $H(T\mathbf{p} \parallel T\mathbf{q}) \leq H(\mathbf{p} \parallel \mathbf{q})$. Moreover, for each $1 \leq k \leq N$, $p_k, q_k > 0$, then $H(T\mathbf{p} \parallel T\mathbf{q}) = H(\mathbf{p} \parallel \mathbf{q})$ if and only if the following conditions hold:*

- (i) $\mathbf{p} = \bigoplus_{k=1}^K \mu_k \mathbf{p}_k \otimes \mathbf{r}_k$ and $\mathbf{q} = \bigoplus_{k=1}^K \nu_k \mathbf{q}_k \otimes \mathbf{r}_k$, where $\mathbf{p}_k, \mathbf{q}_k$ denote the m_k -dimensional probability vectors, and \mathbf{r}_k denotes the n_k -dimensional probability vectors, and $\mu_k, \nu_k \geq 0$, $k = 1, \dots, K$; $\sum_{k=1}^K \mu_k = \sum_{k=1}^K \nu_k = 1$, $\sum_{k=1}^K m_k n_k = N$;
- (ii) $T = \bigoplus_{k=1}^K \pi_k \otimes T_k$, $\pi_k \in \mathbf{S}_{m_k}$ and T_k is $n_k \times n_k$ stochastic matrix for each $k = 1, \dots, K$.

Proof. Let ρ, σ, ρ' and σ' be diagonal matrices with diagonal $\mathbf{p}, \mathbf{q}, T\mathbf{p}$ and $T\mathbf{q}$, respectively. Then it follows from the Lemma 2.2 that there is a stochastic $\Phi \in \mathbf{CP}(\mathbb{C}^n, \mathbb{C}^m)$ such that $\Phi(\rho) = \rho', \Phi(\sigma) = \sigma'$. Note that $H(\mathbf{p}||\mathbf{q}) = S(\rho||\sigma)$ and $H(T\mathbf{p}||T\mathbf{q}) = S(\rho'||\sigma')$, so by Lemma 1 we have $H(T\mathbf{p}||T\mathbf{q}) \leq H(\mathbf{p}||\mathbf{q})$. Moreover, if for each $1 \leq k \leq N$, $p_k, q_k > 0$, then the states $\rho, \sigma, \Phi(\rho) = \rho'$ and $\Phi(\sigma) = \sigma'$ are faithful. If $H(T\mathbf{p}||T\mathbf{q}) = H(\mathbf{p}||\mathbf{q})$, then $S(\rho||\sigma) = S(\Phi(\rho)||\Phi(\sigma))$. Since [6]

$$S(\Phi(\rho)||\Phi(\sigma)) = S(\rho||\sigma)$$

if and only if the following statements hold:

- (1) \mathcal{H} and \mathcal{K} can be decomposed by the form $\mathcal{H} = \bigoplus_{k=1}^K \mathcal{H}_k^L \otimes \mathcal{H}_k^R, \mathcal{K} = \bigoplus_{k=1}^K \mathcal{K}_k^L \otimes \mathcal{K}_k^R$, where $\dim \mathcal{H}_k^L = \dim \mathcal{K}_k^L$.
- (2) If Φ_k is the restriction of Φ to $\mathbf{L}(\mathcal{H}_k^L \otimes \mathcal{H}_k^R)$, then $\Phi_k \in \mathbf{T}(\mathcal{H}_k^L \otimes \mathcal{H}_k^R, \mathcal{K}_k^L \otimes \mathcal{K}_k^R)$ and it can be factorized into the form $\Phi_k = Ad_{U_k} \otimes \Phi_k^R$, where $U_k : \mathcal{H}_k^L \longrightarrow \mathcal{K}_k^L$ is unitary operator and $\Phi_k^R \in \mathbf{T}(\mathcal{H}_k^R, \mathcal{K}_k^R)$ is stochastic, $k = 1, \dots, K$.
- (3) The state ρ decomposes as $\rho = \bigoplus_{k=1}^K p_k \rho_k^L \otimes \omega_k^R, \sigma = \bigoplus_{k=1}^K q_k \sigma_k^L \otimes \omega_k^R$, where all the operators are density operators, and $\{p_k\}_{k=1}^K$ and $\{q_k\}_{k=1}^K$ are probability distributions.

Therefore, it follows that the result can be proved by the above decomposition of Φ . \square

Remark 2.4. In [13], it was showed that $S(\Phi(\rho)) = S(\rho)$ if and only if $\Phi^\dagger \circ \Phi(\rho) = \rho$ while the explicit construction of the state ρ and the quantum operation Φ are given. We can employ the mentioned result to give an explicit construction for T, \mathbf{p} in the identity: $H(T\mathbf{p}) = H(\mathbf{p})$. The proof is trivially and omitted.

3. RELATIVE ENTROPY OF STOCHASTIC MATRICES

In this section, the entropy of stochastic matrices is discussed. For the entropy of stochastic matrices, more details can be found in [12]. We will go deeper within the entropy concerning stochastic matrices and derive some conditions on the (strong) additivity for the stochastic matrices. The notion of relative entropy of stochastic matrices is introduced by mimicking quantum relative entropy. Some properties of this concept are listed and the connection between the entropy of the stochastic quantum operations and that of stochastic matrices are discussed.

To be specific, for any $N \times N$ stochastic matrix $T = [t_{\mu\nu}]$, the *weighted entropy* [12] of T by a probability vector $\mathbf{p} = [p_1, \dots, p_N]^\top$ is defined by $H_{\mathbf{p}}(T) = \sum_{\nu=1}^N p_\nu H(\mathbf{t}_\nu)$, where $T = [\mathbf{t}_1, \dots, \mathbf{t}_N]$ and $\mathbf{t}_\nu = [t_{1\nu}, \dots, t_{N\nu}]^\top$ is the ν th column vector of T . In particular, $H(T) = \frac{1}{N} \sum_{\nu=1}^N H(\mathbf{t}_\nu)$ is defined for $\mathbf{p} = \frac{1}{N}[1, \dots, 1]^\top$.

For any two $N \times N$ stochastic matrices A and B , the *relative entropy* between A and B with respect to a probability vector $\mathbf{p} = [p_1, \dots, p_N]^\top$ is defined by $H_{\mathbf{p}}(A||B) = \sum_{\nu=1}^N p_\nu H(\mathbf{a}_\nu||\mathbf{b}_\nu)$,

where \mathbf{a}_ν and \mathbf{b}_ν are the ν th column vectors of A and B , respectively. Similarly, $H(A\|B) = \frac{1}{N} \sum_{\nu=1}^N H(\mathbf{a}_\nu\|\mathbf{b}_\nu)$ is defined for $\mathbf{p} = \frac{1}{N}[1, \dots, 1]^\top$.

The following conclusions are immediate. That is, $H_{\mathbf{p}}(\cdot)$ is a nonnegative and concave function; $H_{\mathbf{p}}(\cdot\|\cdot)$ is a jointly convex function.

In what follows, the monotonicity of relative entropy of stochastic matrices is obtained.

Theorem 3.1. *If T, A, B are all $N \times N$ stochastic matrices, then*

$$H_{\mathbf{p}}(TA\|TB) \leq H_{\mathbf{p}}(A\|B),$$

where \mathbf{p} is an N -dimensional probability vector. Moreover, if all the components of \mathbf{p} are positive, then

$$H_{\mathbf{p}}(TA\|TB) = H_{\mathbf{p}}(A\|B)$$

if and only if the following conditions hold:

- (i) $\mathbf{a}_j = \bigoplus_{k=1}^K \mu_k^{(j)} \mathbf{p}_k^{(j)} \otimes \mathbf{r}_k$ and $\mathbf{b}_j = \bigoplus_{k=1}^K \nu_k^{(j)} \mathbf{q}_k^{(j)} \otimes \mathbf{r}_k$, where $\mathbf{p}_k^{(j)}, \mathbf{q}_k^{(j)}$ denote m_k -dimensional probability vectors, and \mathbf{r}_k are n_k -dimensional probability vectors and $\forall k : \mu_k, \nu_k \geq 0, \sum_{k=1}^K \mu_k^{(j)} = \sum_{k=1}^K \nu_k^{(j)} = 1$ for each $j = 1, \dots, N, \sum_{k=1}^K m_k n_k = N$;
- (ii) $T = \bigoplus_{k=1}^K \pi_k \otimes T_k$, where $\pi_k \in \mathbf{S}_{m_k}$ and T_k is $n_k \times n_k$ stochastic matrix for each k .

Proof. By the definition of relative entropy for stochastic matrices, it follows that

$$H_p(A\|B) = \sum_{j=1}^N p_j H(\mathbf{a}_j\|\mathbf{b}_j),$$

where $\mathbf{a}_j = [a_{1j}, \dots, a_{Nj}]^\top$ and $\mathbf{b}_j = [b_{1j}, \dots, b_{Nj}]^\top$ are j th columns of A and B , respectively. Now

$$\begin{aligned} H_{\mathbf{p}}(TA\|TB) &= \sum_{j=1}^N p_j H((TA)_j\|(TB)_j) = \sum_{j=1}^N p_j H(T\mathbf{a}_j\|T\mathbf{b}_j) \\ &\leq \sum_{j=1}^N p_j H(\mathbf{a}_j\|\mathbf{b}_j) = H_{\mathbf{p}}(A\|B). \end{aligned}$$

Thus it follows from the above process in the proof that when the components of p are all positive, $H_{\mathbf{p}}(TA\|TB) = H_{\mathbf{p}}(A\|B)$ if and only if $H(T\mathbf{a}_j\|T\mathbf{b}_j) = H(\mathbf{a}_j\|\mathbf{b}_j)$ for each j . By Theorem 2.3, the equality condition can be concluded immediately. \square

Remark 3.2. Now denote $L^{(k)} = [\mathbf{p}_k^{(1)}, \dots, \mathbf{p}_k^{(N)}]$ and $R^{(k)} = [\mathbf{q}_k^{(1)}, \dots, \mathbf{q}_k^{(N)}]$. Let $E^{(k)} = \text{Diag}[\mu_k^{(1)}, \dots, \mu_k^{(N)}]$ and $F^{(k)} = \text{Diag}[\nu_k^{(1)}, \dots, \nu_k^{(N)}]$. The explicit forms of A, B can be written as

$$A = \bigoplus_{k=1}^K E^{(k)} L^{(k)} \otimes \mathbf{r}_k \text{ and } B = \bigoplus_{k=1}^K F^{(k)} R^{(k)} \otimes \mathbf{r}_k,$$

where $L^{(k)}$ and $R^{(k)}$ are any stochastic matrices, where $k = 1, \dots, K$. Furthermore, $\sum_{k=1}^K E^{(k)} = \sum_{k=1}^K F^{(k)} = \mathbb{1}$.

For a finite collection $\{B^{(i)}\}$ of $N \times N$ stochastic matrices, denote $\bar{B} = \sum_i \lambda_i B^{(i)}$, where $\{\lambda_i\}$ is a probability vector. The χ -quantity for $\{B^{(i)}\}$ is defined by $\chi_{\mathbf{p}}(\{\lambda_i, B^{(i)}\}) = \sum_i \lambda_i H_{\mathbf{p}}(B^{(i)} \| \bar{B})$, where \mathbf{p} is a probability vector. It is easily seen from Theorem 3.1 that

- (i) $\chi_{\mathbf{p}}(\{\lambda_i, B^{(i)}\}) = H_{\mathbf{p}}(\sum_i \lambda_i B^{(i)}) - \sum_i \lambda_i H_{\mathbf{p}}(B^{(i)});$
- (ii) $\sum_i \lambda_i H_{\mathbf{p}}(B^{(i)} \| D) = \chi_{\mathbf{p}}(\{\lambda_i, B^{(i)}\}) + H_{\mathbf{p}}(\bar{B} \| D)$, i.e., $\sum_i \lambda_i H_{\mathbf{p}}(B^{(i)} \| D) = \sum_i \lambda_i H_{\mathbf{p}}(B^{(i)} \| \bar{B}) + H_{\mathbf{p}}(\bar{B} \| D)$, where D is $N \times N$ stochastic matrix and p is an N -dimensional probability vector.
- (iii) Assume that T is $N \times N$ stochastic matrix. Then $\chi_{\mathbf{p}}(\{\lambda_i, TB^{(i)}\}) \leq \chi_{\mathbf{p}}(\{\lambda_i, B^{(i)}\})$ if and only if $H_{\mathbf{p}}(TB) - H_{\mathbf{p}}(B)$ is a convex function in its argument stochastic matrix B ; moreover, $\chi_{\mathbf{p}}(\{\lambda_i, TB^{(i)}\}) \leq \chi_{\mathbf{p}}(\{\lambda_i, B^{(i)}\})$.

In [12], W. Słomczyński obtained that given any $N \times N$ stochastic matrices X, Y, Z for which \mathbf{p} is their common invariant probability vector, i.e. $X\mathbf{p} = Y\mathbf{p} = Z\mathbf{p} = \mathbf{p}$. Then :

- (i) $H_{\mathbf{p}}(Y) \leq H_{\mathbf{p}}(XY) \leq H_{\mathbf{p}}(X) + H_{\mathbf{p}}(Y);$
- (ii) $H_{\mathbf{p}}(XYZ) + H_{\mathbf{p}}(Y) \leq H_{\mathbf{p}}(XY) + H_{\mathbf{p}}(YZ).$

The following result is to deal with the saturation of the above two inequalities.

Proposition 3.3. (i) If $T \in \mathbf{B}_N$, A is $N \times N$ stochastic matrix and \mathbf{p} is an N -dimensional probability vector with all positive components, then $H_{\mathbf{p}}(TA) = H_{\mathbf{p}}(A)$ if and only if $T^{\top}TA = A$;

(ii) If $X = X_L \otimes \pi_R$ and $Y = \pi_L \otimes Y_R$ for X_L being stochastic matrix of size $m \times m$, $\pi_L \in \mathbf{S}_m$, Y_R being stochastic matrix of size $n \times n$, $\pi_R \in \mathbf{S}_n$, then $H(XY) = H(X) + H(Y)$;

(iii) If $X = \bigoplus_{k=1}^K X_k^L \otimes \pi_k^R$, $Y = \bigoplus_{k=1}^K Y_k^L \otimes Y_k^R$ and $Z = \bigoplus_{k=1}^K \pi_k^L \otimes Z_k^R$ for X_k^L being stochastic matrix of size $m_k \times m_k$, $\pi_k^L \in \mathbf{S}_{m_k}$, Y_k^R being stochastic matrix of size $n_k \times n_k$, $\pi_k^R \in \mathbf{S}_{n_k}$, then $H(XYZ) + H(Y) = H(XY) + H(YZ)$.

Proof. (i) Since each component p_j are positive, it follows that $H_{\mathbf{p}}(TA) = H_{\mathbf{p}}(A)$ if and only if $H(T\mathbf{a}_j) = H(\mathbf{a}_j)$ for every j , where all \mathbf{a}_j 's are the j th column vector of A . By the result in [9, 13], i.e. for $B \in \mathbf{B}_N$, $H(B\mathbf{p}) = H(\mathbf{p})$ if and only if $B^{\top}B\mathbf{p} = \mathbf{p}$, we get that $H(T\mathbf{a}_j) = H(\mathbf{a}_j)$ for every j if and only if $T^{\top}T\mathbf{a}_j = \mathbf{a}_j$ for all j ; that is, the proof is concluded.

(ii) Since $XY = X_L \pi_L \otimes \pi_R Y_R$, it follows that $H(XY) = H(X_L \pi_L \otimes \pi_R Y_R) = H(X_L) + H(Y_R)$, which implies that the conclusion.

(iii) Since

$$XYZ = \bigoplus_{k=1}^K X_k^L Y_k^L \pi_k^L \otimes \pi_k^R Y_k^R Z_k^R,$$

it follows that

$$\begin{aligned} H(XYZ) &= \sum_k \lambda_k [H(X_k^L Y_k^L) + H(Y_k^R Z_k^R)], \\ H(XY) &= \sum_k \lambda_k [H(X_k^L Y_k^L) + H(Y_k^R)], \\ H(YZ) &= \sum_k \lambda_k [H(Y_k^L) + H(Y_k^R Z_k^R)], \\ H(Y) &= \sum_k \lambda_k [H(Y_k^L) + H(Y_k^R)], \end{aligned}$$

where $\lambda_k = m_k n_k / N$ and $\sum_k m_k n_k = N$. Combining all these expressions gives the desired result. \square

4. THE RELATIONSHIP BETWEEN QUANTUM OPERATIONS AND BI-STOCHASTIC MATRICES

For any CP super-operators Φ and Ψ , with corresponding their Kraus representations: $\Phi = \sum_i \mathbf{A} \mathbf{d}_{M_i}$ and $\Psi = \sum_j \mathbf{A} \mathbf{d}_{N_j}$, respectively. It is easily seen that $\Phi \otimes \Psi = \sum_{i,j} \mathbf{A} \mathbf{d}_{M_i \otimes N_j}$, and $J(\Phi \otimes \Psi) = J(\Phi) \otimes J(\Psi)$; denote $\Psi^\top = \sum_j \mathbf{A} \mathbf{d}_{N_j}^\top$. Then $J(\Phi \circ \Psi) = \Phi \otimes \mathbb{1}(J(\Psi)) = \mathbb{1} \otimes \Psi^\top(J(\Phi)) = \Phi \otimes \Psi^\top(\text{vec}(\mathbb{1}) \text{vec}(\mathbb{1})^\dagger)$.

Let $\Phi, \Psi \in \mathbf{CP}(\mathcal{H})$ be stochastic, the *relative entropy* between Φ and Ψ is defined by $\mathbf{S}(\Phi \| \Psi) = \mathbf{S}(\rho(\Phi) \| \rho(\Psi))$. If $\Lambda \in \mathbf{CP}(\mathcal{H})$ is also bi-stochastic, then $\mathbf{S}(\Lambda \circ \Phi \| \Lambda \circ \Psi) \leq \mathbf{S}(\Phi \| \Psi)$. This can be seen easily from the Lemma 2.1. Indeed, $\mathbf{S}(\Lambda \otimes \mathbb{1}_{\mathbf{L}(\mathcal{H})}(\rho(\Phi)) \| \Lambda \otimes \mathbb{1}_{\mathbf{L}(\mathcal{H})}(\rho(\Psi))) \leq \mathbf{S}(\rho(\Phi) \| \rho(\Psi)) = \mathbf{S}(\Phi \| \Psi)$ since $\Lambda \otimes \mathbb{1}_{\mathbf{L}(\mathcal{H})}$ is bi-stochastic whenever Λ is bi-stochastic.

Assume that Φ is a CP stochastic super-operator for which the Kraus decomposition can be written as $\Phi = \sum_\mu \mathbf{A} \mathbf{d}_{T_\mu}$. Define *Kraus matrix* [1] for Φ as $B(\Phi) := \sum_\mu T_\mu \bullet T_\mu^*$, where \bullet denotes Shur product of matrices and $*$ means that entry-wise complex conjugate of a matrix. Hence the (i, j) th entry b_{ij} of B can be described by $b_{ij} = \sum_\mu t_{ij}^\mu \overline{t_{ij}^\mu}$, where $T_\mu = [t_{ij}^\mu]$ and the bar means the complex conjugate of complex numbers.

For any two Hermitian matrices X and Y , X is *majorized* by Y , denoted by $X < Y$, if there is a CP bi-stochastic super-operator Φ such that $X = \Phi(Y)$. The well-known Shur's theorem states that $\text{Diag}(X) < X$ for any square matrix X , see [1]. Thus for any bi-stochastic quantum operation Λ , it follows that $\Lambda(\rho) < \rho$. Moreover, $J(\Phi \circ \Psi) < J(\Phi)$ and $J(\Phi \circ \Psi) < J(\Psi)$ for any two bi-stochastic quantum operations Φ and Ψ .

In what follows, some properties of Kraus matrix is listed below.

- Proposition 4.1.** (i) For a given (bi-)stochastic super-operator $\Phi \in \mathbf{CP}(\mathcal{H})$, $B(\Phi)$ is a (bi-)stochastic matrix.
- (ii) $B(\Phi)$ is well-defined, i.e., it is independent of the different Kraus decompositions for Φ and just depends on Φ itself.

(iii) $B(\Phi)$ is a convex function with respect to argument Φ , i.e.,

$$B(t\Phi_1 + (1-t)\Phi_2) = tB(\Phi_1) + (1-t)B(\Phi_2) \text{ for any } \Phi_1, \Phi_2 \text{ and all } t \in [0, 1].$$

(iv) Denote $\mathcal{M}(B) = \{\Phi | \Phi \in \mathcal{T}(\mathcal{H}) \text{ is CP stochastic super-operator and } B(\Phi) = B\}$. Then $\mathcal{M}(B)$ is a nonempty convex set.

(v) $B(\Theta_1 \otimes \Theta_2) = B(\Theta_1) \otimes B(\Theta_2)$ for any stochastic quantum operations Θ_1 and Θ_2 .

(vi) Assume that \mathcal{H} and \mathcal{K} are M and N dimensional Hilbert spaces, respectively. If $\Lambda \in \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ is CP and stochastic and can be described by $\Lambda = \sum_k \lambda_k \Phi_k \otimes \Psi_k$, where $\{\Phi_k\} \in \mathcal{T}(\mathcal{H})$ and $\{\Psi_k\} \in \mathcal{T}(\mathcal{K})$ are two collections of CP stochastic super-operators, then $B(\Lambda) = \sum_k \lambda_k B(\Phi_k) \otimes B(\Psi_k)$, where $\lambda = \{\lambda_k\}_k$ is a finite probability vector.

Proof. (i) The proof is trivially.

(ii) Assume that $\Phi = \sum_{\mu=1}^{N^2} Ad_{E_\mu} = \sum_{\nu=1}^{N^2} Ad_{F_\nu}$. By the unitary freedom of quantum operations, there is a $N^2 \times N^2$ unitary matrix $U = [u_{\mu\nu}]$ such that $E_\mu = \sum_{\nu=1}^{N^2} u_{\mu\nu} F_\nu$. Then

$$\begin{aligned} \sum_{\mu=1}^{N^2} E_\mu \bullet E_\mu^* &= \sum_{\mu=1}^{N^2} \left(\sum_{\nu=1}^{N^2} u_{\mu\nu} F_\nu \right) \bullet \left(\sum_{\kappa=1}^{N^2} u_{\mu\kappa} F_\kappa \right)^* = \sum_{\nu,\kappa=1}^{N^2} \left(\sum_{\mu=1}^{N^2} u_{\mu\nu} \bar{u}_{\mu\kappa} \right) F_\nu \bullet F_\kappa^* \\ &= \sum_{\nu,\kappa=1}^{N^2} \delta_{\nu\kappa} F_\nu \bullet F_\kappa^* = \sum_{\nu=1}^{N^2} F_\nu \bullet F_\nu^*, \end{aligned}$$

which implies that $B(\Phi)$ is well-defined.

(iii) Choose any two stochastic quantum operations Φ_1 and Φ_2 with their corresponding Kraus decomposition: $\Phi_1 = \sum_\mu Ad_{S_\mu}$ and $\Phi_2 = \sum_\nu Ad_{T_\nu}$. Let $t \in [0, 1]$. Then Kraus decomposition for $t\Phi_1 + (1-t)\Phi_2$ is $t\Phi_1 + (1-t)\Phi_2 = t \sum_\mu Ad_{S_\mu} + (1-t) \sum_\nu Ad_{T_\nu}$, which implies that the Kraus matrix for $t\Phi_1 + (1-t)\Phi_2$ is

$$\begin{aligned} B(t\Phi_1 + (1-t)\Phi_2) &= \sum_\mu (\sqrt{t} S_\mu) \bullet (\sqrt{t} S_\mu^*) + \sum_\nu (\sqrt{1-t} T_\nu) \bullet (\sqrt{1-t} T_\nu^*) \\ &= t \sum_\mu S_\mu \bullet S_\mu^* + (1-t) \sum_\nu T_\nu \bullet T_\nu^* \\ &= tB(\Phi_1) + (1-t)B(\Phi_2). \end{aligned}$$

(iv) If $\Psi_1, \Psi_2 \in \mathcal{M}(B)$, it follows from the result of (iii) that $B(t\Psi_1 + (1-t)\Psi_2) = tB(\Psi_1) + (1-t)B(\Psi_2) = B$ since $B(\Psi_1) = B(\Psi_2) = B$, which implies $t\Psi_1 + (1-t)\Psi_2 \in \mathcal{M}(B)$. The fact that $\mathcal{M}(B)$ is not empty is clearly.

(v) Let the Kraus decompositions for Θ_1 and Θ_2 are $\Theta_1 = \sum_m Ad_{S_m}$ and $\Theta_2 = \sum_\mu Ad_{T_\mu}$. Then $\Theta_1 \otimes \Theta_2 = \sum_{m,\mu} Ad_{S_m \otimes T_\mu}$. Now

$$\begin{aligned} B(\Theta_1 \otimes \Theta_2) &= \sum_{m,\mu} (S_m \otimes T_\mu) \bullet (S_m^* \otimes T_\mu^*) = \left(\sum_m S_m \bullet S_m^* \right) \otimes \left(\sum_\mu T_\mu \bullet T_\mu^* \right) \\ &= B(\Theta_1) \otimes B(\Theta_2). \end{aligned}$$

(vi) It follows trivially from combining the above conclusions (iii) and (v).

□

Remark 4.2. In the above Proposition 4.1(iv), it is known that $\mathcal{M}(B)$ is a nonempty convex set. In fact, it is also compact. Thus the question naturally arises: what is the extreme points of $\mathcal{M}(B)$? Note that in [8], Parthasarathy gave a characterization of extremal quantum states of composite systems with fixed marginal states. Subsequently, Rudolph gave another characterization about it in [11]. Therefore, our question can be described in terms of the language as in [8, 11] under the additional condition that the diagonals of Jamiołkowski state is fixed.

Remark 4.3. Assume that \mathcal{H} is an N -dimensional Hilbert space and $\Phi, \Psi \in \text{CP}(\mathcal{H})$ for which

$$J(\Phi) = \sum_{m,\mu=1}^N p_{m\mu} |m\mu\rangle\langle m\mu| \text{ and } J(\Psi) = \sum_{m,\mu=1}^N q_{m\mu} |m\mu\rangle\langle m\mu|.$$

By the stochasticity, $\sum_{\mu=1}^N p_{m\mu} = 1$ and $\sum_{m=1}^N p_{m\mu} = 1$; $\sum_{\mu=1}^N q_{m\mu} = 1$ and $\sum_{m=1}^N q_{m\mu} = 1$. Then $J(\Phi \circ \Psi) = \sum_{m,\mu=1}^N [B(\Phi)B(\Psi)]_{m\mu} |m\mu\rangle\langle m\mu|$, where $B(\Phi) = [p_{m\mu}]$ and $B(\Psi) = [q_{m\mu}]$, which implies that

$$B(\Phi \circ \Psi) = B(\Phi)B(\Psi), B(\Psi \circ \Phi) = B(\Psi)B(\Phi)$$

and

$$\mathcal{S}(\Phi \circ \Psi) = H(B(\Phi \circ \Psi)) + \log N, \mathcal{S}(\Psi \circ \Phi) = H(B(\Psi \circ \Phi)) + \log N.$$

Now $\mathcal{S}^{\text{map}}(\Phi) + \mathcal{S}^{\text{map}}(\Psi) - \mathcal{S}^{\text{map}}(\Phi \circ \Psi) = H(B(\Phi)) + H(B(\Psi)) - H(B(\Psi)B(\Phi)) + \log N$.

Generally speaking, $B(\Phi \circ \Psi) \neq B(\Phi)B(\Psi)$ for two stochastic super-operators $\Phi, \Psi \in \text{CP}(\mathcal{H})$. This fact shows that if both $J(\Phi)$ and $J(\Psi)$ are diagonal, then $B(\Phi \circ \Psi) = B(\Phi)B(\Psi)$. There is a question which can be formulated as follows: what is a sufficient and necessary condition for $B(\Phi \circ \Psi) = B(\Phi)B(\Psi)$ for stochastic super-operators $\Phi, \Psi \in \text{CP}(\mathcal{H})$. It is *conjectured* that $H_{\mathbf{p}}(B(\Phi \circ \Psi)) \leq H_{\mathbf{p}}(B(\Phi)B(\Psi))$ for any stochastic super-operators $\Phi, \Psi \in \text{CP}(\mathcal{H})$, where \mathbf{p} is any N -dimensional probability vector.

Proposition 4.4. Assume that \mathcal{H} is a N -dimensional Hilbert space.

- (i) If $\Phi \in \text{CP}(\mathcal{H})$ is stochastic, then: $\mathcal{S}^{\text{map}}(\Phi) \leq H(B(\Phi)) + \log N$;
- (ii) If $\Phi, \Psi \in \text{CP}(\mathcal{H})$ is stochastic, then: $H(B(\Phi)B(\Psi)) \leq \mathcal{S}(\Phi \parallel \Psi)$;

Proof. (i) By Shur's lemma, it follows that $\text{Diag}(J(\Phi)) \prec J(\Phi)$ which is equivalent to $\text{Diag}(\rho(\Phi)) \prec \rho(\Phi)$. Since $\langle m|B(\Phi)|\mu\rangle = \langle m\mu|J(\Phi)|m\mu\rangle$, it can be seen that

$$\mathcal{S}^{\text{map}}(\Phi) = \mathcal{S}(\rho(\Phi)) \leq \mathcal{S}(\text{Diag}(\rho(\Phi))) = H(B(\Phi)) + \log N.$$

Furthermore, $\mathcal{S}^{\text{map}}(\Phi) = H(B(\Phi)) + \log N$ when Φ is represented by a diagonal dynamical matrix $J(\Phi)$.

- (ii) There exists a CP bi-stochastic super-operator Λ such that $\Lambda(\rho) = \text{Diag}(\rho)$ since $\text{Diag}(\rho) \prec \rho$ which follows from Shur's lemma. Thus it follows from Lemma 2.1 that

$$\begin{aligned} S(\Phi\|\Psi) &= S(\rho(\Phi)\|\rho(\Psi)) \geq S(\text{Diag}[\rho(\Phi)]\|\text{Diag}[\rho(\Psi)]) \\ &= \frac{1}{N} \sum_{j=1}^N H(B(\Phi)_j\|B(\Psi)_j) = H(B(\Phi)\|B(\Psi)). \end{aligned}$$

□

Remark 4.5. Let $\widetilde{s}(\Phi) = H(B(\Phi)) - S^{\text{map}}(\Phi)$ for stochastic super-operator $\Phi \in \text{CP}(\mathcal{H})$. Then for a collection $\{\Phi_k\}$ of stochastic super-operator in $\text{CP}(\mathcal{H})$ such that $\Phi = \sum_k \lambda_k \Phi_k$, $\chi(\{\lambda_k, B(\Phi_k)\}) \leq \chi(\{\lambda_k, \Phi_k\})$ if and only if $\widetilde{s}(\sum_k \lambda_k \Phi_k) \leq \sum_k \lambda_k \widetilde{s}(\Phi_k)$; i.e., $\widetilde{s}(\Phi)$ is a convex function in its argument Φ .

REFERENCES

- [1] I. Bengtsson and K. Życzkowski, Geometry of quantum states, Cambridge University Press, pp. 315(2006).
- [2] M. D. Choi, Completely positive linear maps on complex matrices, Linear algebra and its applications, **10**:285-290(1975).
- [3] K. Kraus, States, effects, and operations : fundamental notions of quantum theory, Publisher: Springer-Verlag (1983).
- [4] C. K. Li and Y. T. Poon, Interpolation problems by completely positive maps, arXiv: math.FA/1012.1675 (2010).
- [5] G. Lindblad, Completely Positive Maps and Entropy Inequalities, Commun. math. Phys. **40**: 147–151(1975).
- [6] M. Mosonyi and D. Petz: Structure of Sufficient Quantum Coarse Grainings, Lett. Math. Phys. **68**: 19–30(2004).
- [7] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge (2000).
- [8] K.R. Parthasarathy, Extremal quantum states in coupled systems, Ann. I. H. Poincaré–PR **41**:257–268(2005).
- [9] A. B. Poritz and J. A. Poritz: On entropy-preserving stochastic averages, Linear Algebra and its Applications **434**:1425–1442(2011).
- [10] W. Roga, M. Fannes and K. Życzkowski, Composition of quantum states and dynamical subadditivity. J. Phys. A: Math. Theor. **41**, 035305 (2008).
- [11] O. Rudolph, On extremal quantum states of composite systems with fixed margins, J. Math. Phys. **45**:4035–4041(2004).
- [12] W. Słomczyński, Subadditivity of entropy for stochastic matrices, Open& Information Dyn. **9**: 201–206(2002).
- [13] L. Zhang and J. Wu, Von Neumann Entropy-Preserving Quantum Operations, Phys. Lett. A **375**(47): 4163–4165(2011).

DEPARTMENT OF MATHEMATICS, ZHEJIANG UNIVERSITY, HANGZHOU, PEOPLE'S REPUBLIC OF CHINA

E-mail address: godyalin@163.com, linyz@zju.edu.cn